# Monte Carlo Sampling Methods Homework 5

Zihan (Steven) Zhang

December 14, 2022

## 1 Exercise 64

In this question, we first define the XY model of statistical physics. Consider a set of 2-dimensional vectors, $\vec{\sigma}_i \in \mathbb{R}^2$ indexed by the 1-dimensional periodic lattice $\mathbb{Z}_L$ and with $||\vec{\sigma}_i||_2 = 1$. Thus, the density:

$$\pi(\vec{\sigma}) = \frac{e^{\beta} \sum_{i \leftrightarrow j} \vec{\sigma}_i \vec{\sigma}_j}{\mathcal{Z}}. \tag{1}$$

In terms of the angles $\theta_i \in [-\pi, \pi)$ of the vectors $\vec{\sigma}_i$, we could rewrite the density as:

$$\pi(\theta) = \frac{e^{\beta} \sum_{i \leftrightarrow j} \cos(\theta_i - \theta_j)}{\mathcal{Z}}. \tag{2}$$

Thus, implementing the simple recursion formula:

$$X_h^{(k+1)} = X_h^{(k)} + hb\left(X_h^{(k)}\right) + \sqrt{h}\sigma\left(X_h^{(k)}\right)\xi^{(k+1)}. \tag{3}$$

Here each $\xi^{(k)}$ is a vector of independent random variables with $\mathbf{P}[(\xi^{(k)})_i = 1] = \mathbf{P}[(\xi^{(k)})_i = -1] = 1/2$ or a vector of independent standard Gaussian random variables that has a limiting generator:

$$\mathcal{L}f(x) = \nabla f(x) b(x) + \frac{1}{2}\operatorname{trace}\left(\sigma(x)\sigma^{\mathrm{T}}(x) D^2 f(x)\right). \tag{4}$$

We could have the discrete Markov chain as:

$$X_h^{(k+1)} = X_h^{(k)} - h(S)\left(X_h^{(k)}\right)\nabla^{\mathrm{T}} H\left(X_h^{(k)}\right) + h\operatorname{div}(S)\left(X_h^{(k)}\right) + \sqrt{2hS\left(X_h^{(k)}\right)}\xi^{(k)}. \tag{5}$$

for $S$ the identity matrix and $\mathbf{E}[\xi^{(k)}] = 0$ and $\mathbf{cov}[\xi^{(k)}] = I$ (and finite higher moments). It could be simplified as:

$$X^{(k+1)} = X^{(k)} + h\nabla^T \log\pi\left(X^k\right) + \sqrt{2h}\xi^{(k+1)}. \tag{6}$$

We should write a routine to sample the XY model using Equation (3) and a Metropolized version. The total number of time steps is scaled like $1/h$. Also, we should make the comparison in terms of the integrated autocorrelated time of the variable:

$$\frac{M_1(\sigma)}{||M(\sigma)||_2}. \tag{7}$$

which is the cosine of the angle of the magnetization vector:

$$M(\sigma) = \sum_{i=0}^{L-1} \vec{\sigma}_i \in \mathbb{R}^2. \tag{8}$$

Firstly we fix $L = 10$ and $\beta = 0.1$ and change different $h$ (equispaced variation from $10^{-2.5}$ to $10^{-1}$) to calculate IAT for (un)metropolized schemes. Figure 1 shows the results with the inversely proportional reference line: when $h$ increases, the IAT for both the Metropolized version and the standard overdamped Langevin schemes will decrease. Also, we fix $h = 10^{-1}$ and slightly change $L$. Figure 2 shows the result for $h = 10^{-1}$ and $\beta = 0.1$. The simulation process is quite stochastic, and the relation between $L$ and IAT is unclear. If the perturbation is small, it is confident to say that IAT is decreasing when $L$ increases.

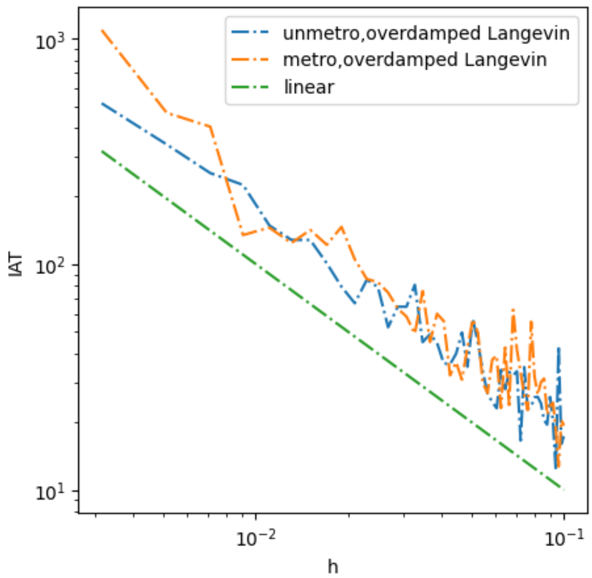


**Figure 1:** Exercise 64: Relationship between $h$ and IAT.

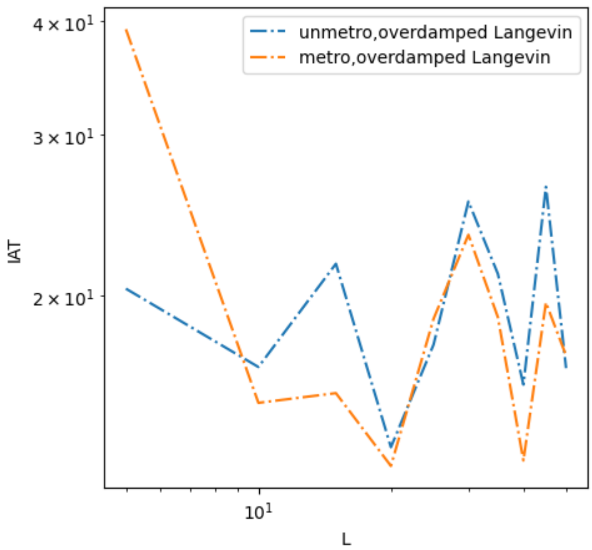


**Figure 2:** Exercise 64: Relationship between $L$ and IAT.

## 2 Exercise 65

In this question, we could write a routine to sample the XY model using Hybrid Monte Carlo model. Specifically, we would independently sample variable $\tilde{Y}(k)$ from the density proportional to $e^{-K(\tilde{x})}$. Set $Y^{(k+1)} = y_h^{(n)}\left(\hat{X}^{(k)}, \tilde{Y}^{(k)}\right)$ where $y_h^{(\ell)}$ solves the Velocity Verlet scheme discretization with initial conditions $x$ and $n$ is chosen by the user. Also, with the probability:

$$p_{acc}\left(X^{(k)}, Y^{(k+1)}\right) = \min\left\{1, \frac{\pi_H\left(Y^{(k+1)}\right)}{\pi_H\left(\hat{X}^{(k)}, \tilde{Y}^{(k)}\right)}\right\}. \tag{9}$$

we set $X^{(k+1)} = Y^{(k+1)}$. Otherwise we set $X^{(k+1)} = X^{(k)}$. Here we choose $K(y) = \frac{||y||_2^2}{2}$, $\hat{J} = I$, $n = 5$, $L = 10$, and $\beta = 0.1$. Similarly to Section 1, Figure 3 describes the relationship of $h$ and IAT. The IAT for (un)metropolized schemes would decrease when $h$ increases.

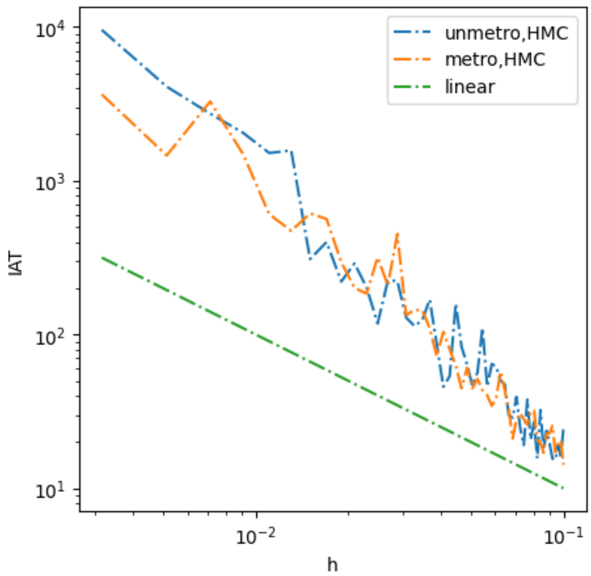


**Figure 3:** Exercise 65: Relationship bewteen $h$ and IAT.

# 3 Exercise 66

In this question, we could sample the XY model following:

$$\tilde{X}_h^{(\ell+1)} = \tilde{X}_h'' + \frac{h}{2}\hat{J}^{\mathrm{T}}\nabla^{\mathrm{T}}\log\pi\left(\hat{X}_h^{(\ell+1)}\right). \tag{10}$$

where $\xi$ is the sequence of independent $\tilde{d}$ dimensional Gaussian random vectors with mean zero and identity covariance (and finite higher moments). Let $Y^{(k+1)}$ be the result of a single step of Equation (10) starting from the initial point $X^{(k)} = \left(\hat{X}^{(k)}, \tilde{X}^{(k)}\right)$. With probability:

$$p_{\text{acc}}\left(X^{(k)}, Y^{(k+1)}\right) = \min\left\{1, \frac{\pi_H\left(Y^{(k+1)}\right)r\left(\hat{Y}^{(k+1)}, -\tilde{Y}^{(k+1)}, \hat{X}^{(k)}, -\tilde{X}^{(k)}\right)}{\pi_H\left(X^{(k)}\right)r\left(X^{(k)}, Y^{(k+1)}\right)}\right\} \tag{11}$$

set $X^{(k+1)} = Y^{(k+1)}$. Otherwise set $X^{(k+1)} = \left(\hat{X}^{(k)}, -\tilde{X}^{(k)}\right)$. The parameter setting $(K, J, n, h, L, \beta)$ is similar to Section 2. Figure 4 describes the relationship between $h$ and IAT. Figure 5 describes the relationship between $\gamma$ and IAT. By comparing the results of IAT values to Section 1 and 2, we conclude the Exercise 66 scheme (using Equation (10) with $\hat{J} = \hat{I}$) is the preferred one for sampling the XY model.

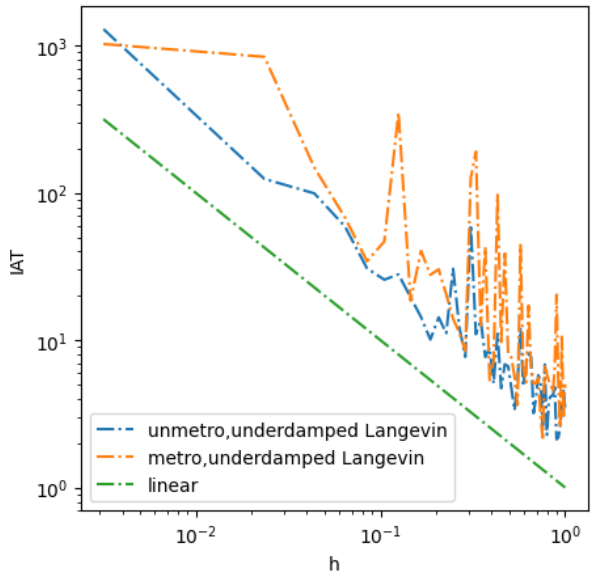


**Figure 4:** Exercise 66: Relationship between $h$ and IAT.

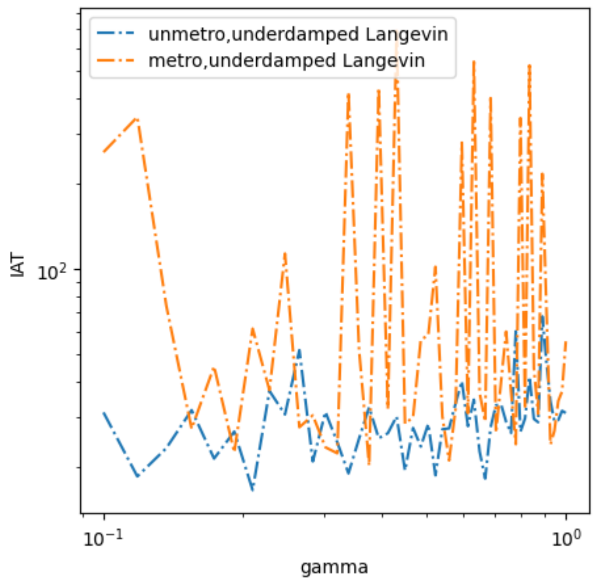


**Figure 5:** Exercise 66: Relationship between $\gamma$ and IAT.